

Sistem *Online Content Based Image Retrieval* menggunakan Identifikasi Dominan Warna pada Foreground Objek

Nana Ramadijanti¹, Setiawardhana¹, Fitria Purnamasari²

Dosen¹, Mahasiswa²

Politeknik Elektronika Negeri Surabaya

Kampus PENS-ITS Keputih Sukolilo Surabaya 60111

Telp (+62)31-5947280, 5946114, Fax. (+62)31-5946114

Email : nana@eepis-its.edu¹, setia@eepis-its.edu¹, v3a88@yahoo.com²

Abstrak

Content Based Image Retrieval atau CBIR merupakan salah satu bentuk aplikasi computer vision untuk pencarian citra berdasarkan fitur yang ada pada citra itu sendiri. Sistem dibangun dengan memanfaatkan fitur utama dari citra yaitu fitur warna. Ekstraksi fitur warna yang umum dipakai adalah histogram warna, tetapi pada umumnya citra hadir dalam kondisi background yang sangat beragam, hal ini berpengaruh terhadap kebenaran hasil retrieval. Optimasi CBIR yang dilakukan pada penelitian ini adalah menggunakan fitur warna dari *foreground* citra. Pada penelitian ini aplikasi CBIR dibangun dengan input citra query secara *online* dengan menggunakan kamera web.

Proses pencarian dilakukan dengan pengambilan gambar online kemudian dengan metode proses ekstraksi fitur warna sehingga dapat diketahui informasi ciri dari gambar. Pemrosesan awal yang dilakukan adalah normalisasi ukuran (resize 320 x 240), normalisasi cahaya, menghilangkan *background*. Model warna yang digunakan pada ekstraksi fitur adalah HSV. Proses selanjutnya adalah histogram interseksi pada warna HSV.

Hasil pengujian CBIR Online pada sepuluh kategori obyek diperoleh rata-rata kemiripan 94.4%. Histogram interseksi HSV dapat digunakan untuk memperoleh fitur ciri general citra. Pengujian CBIR Online dengan data citra di luar data training rata-rata kemiripan sebesar 90%.

Kata Kunci : *Content Based Image Retrieval*, Histogram HSV, Citra *Foreground*.

1. Pendahuluan

Penelitian *image retrieval* pada umumnya dilakukan dengan menggunakan pendekatan indeks dan informasi citra yang berbasis teks. Teknik pencarian berbasis teks menjadi tidak praktis karena ukuran basis data gambar yang besar dan bersifat subyektif dalam mengartikan gambar dengan teks. Kata kunci yang dikodekan terbatas pada beberapa istilah yang dihasilkan untuk masing-masing referensi gambar. Lebih lanjut beberapa gambar yang dihasilkan akan tampak sangat berbeda dibandingkan dengan keinginan user dari otomatisasi pencarian menggunakan kata kunci.

Teknik tersebut diperbaiki dengan cara *image retrieval* berbasis pendekatan alternative yaitu teknik mencari gambar hanya berdasarkan informasi yang ada pada gambar tersebut. Teknik *image retrieval* yang dipilih disamping dapat mencapai rata-rata kemampuan retrieval yang tinggi, seringkali memberikan konsekuensi waktu komputasi yang tinggi dikarenakan harus memproses dimensi data gambar yang besar.

Ekstraksi fitur warna yang umum dipakai adalah histogram warna, tetapi pada umumnya citra hadir dalam kondisi *background* yang sangat beragam, hal ini berpengaruh terhadap kebenaran hasil *retrieval*. Untuk

mengoptimalkan fitur warna dari objek gambar maka pada penelitian ini fitur warna yang dipakai hanya warna *foreground* objek gambar. Proses untuk mendapatkan ekstraksi fitur warna *foreground* menggunakan interseksi histogram dengan distribusi warna yang digunakan sebanyak 48 warna. Metode yang dibangun adalah dapat meningkatkan kebenaran hasil *retrieval* dari masukan *query* gambar secara *online*. Permasalahan utama pada penelitian ini adalah :

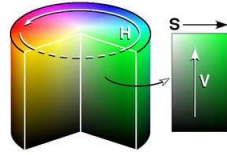
1. Teknik pencarian gambar dengan menggunakan fitur warna
2. Teknik mendapatkan fitur warna yang dominan pada gambar.
3. Pengembangan sistem online kamera dengan performansi dan kualitas yang baik.

Tujuan penelitian ini adalah membangun sistem online CBIR menggunakan fitur warna dominan pada obyek gambar, dengan kontribusi yaitu alternatif pada masukan query sistem CBIR dan metode untuk mencari fitur warna dominan pada objek gambar.

2. Kajian Pustaka

2.1. Model Warna HSV

Model warna HSV mendefinisikan warna dalam terminologi Hue, Saturation dan Value. Hue menyatakan warna sebenarnya, seperti merah, violet, dan kuning. Model ruang warna ini pernah digunakan pada penelitian sebelumnya[3,4,5]. Hue digunakan untuk membedakan warna-warna dan menentukan kemerahan (redness), kehijauan (greeness), dsb, dari cahaya. Hue berasosiasi dengan panjang gelombang cahaya. Saturation menyatakan tingkat kemurnian suatu warna, yaitu mengindikasikan seberapa banyak warna putih diberikan pada warna. Value adalah atribut yang menyatakan banyaknya cahaya yang diterima oleh mata tanpa memperdulikan warna.



Gambar 1. Model Warna HSV.

Karena model warna HSV merupakan model warna yang diturunkan dari model warna RGB maka untuk mendapatkan warna HSV ini, kita harus melakukan proses konversi warna dari RGB ke HSV. HSV merupakan salah satu cara untuk mendefinisikan warna yang didasarkan pada roda warna. Hue mengukur sudut sekitar roda warna (merah pada 0 derajat, 120 derajat di hijau, biru, di 240 derajat). Saturation yang menunjukkan pada radius roda warna sehingga menunjukkan proporsi antara gelap (pusat) untuk warna ke putih murni (di luar). Value Menunjukkan nilai kecerahan. Hue memiliki nilai antara 0 hingga 360 (derajat), Saturation and Value berkisar dari 0 hingga 100%.

2.2. Histogram Interseksi

Histogram interseksi pernah dilakukan oleh peneliti sebelumnya[3,4], dengan melakukan perhitungan jarak dengan membandingkan dua histogram $h1$ dan $h2$ dan terhadap n bin dengan mengambil nilai interseksi dari kedua histogram tersebut.

$$D_H(h1, h2) = 1 - \frac{\sum_{i=1}^n \min(h1_i, h2_i)}{\sum_{i=1}^n h1_i}$$

Normalisasi diperlukan karena pada saat menggunakan image dengan ukuran yang berbeda, fungsi jarak ini bukanlah sebuah matrik yang seharusnya $DH(g,h) \neq DH(h,g)$. Supaya menjadi sebuah matrik yang valid, pada histogram diperlukan :

$$h1^n = \frac{h1}{\sum_{i=1}^n h1_i}$$

Untuk histogram yang ternormalisasi (jumlah total sama dengan 1), interseksi histogram adalah :

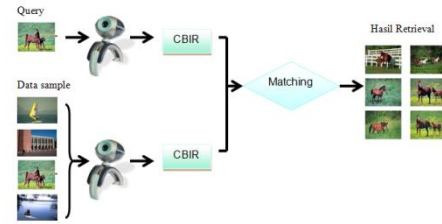
$$D_H^n(h1^n, h2^n) = 1 - \sum_{i=1}^n |h1_i^n - h2_i^n|$$

Persamaan diatas adalah model matriks jarak Minkowski dengan $k=1$. Sifat interseksi histogram dapat menghilangkan bagian tertentu (occlusion), dimana apabila sebuah objek dalam suatu image dihilangkan pada bagian tersebut, bagian yang kelihatan masih mempunyai kontribusi untuk kesamaan atau similaritas.

3. Perancangan Sistem

3.1. Perancangan Perangkat Lunak

Sistem secara keseluruhan dirancang dengan dengan dimulai dari masukan data *query*, kemudian dilakukan proses pemilahan *background* dan *foreground*, dilakukan interseksi histogram HSV, kemudian pencocokan dan hasil *retrieval* ditampilkan di layar.

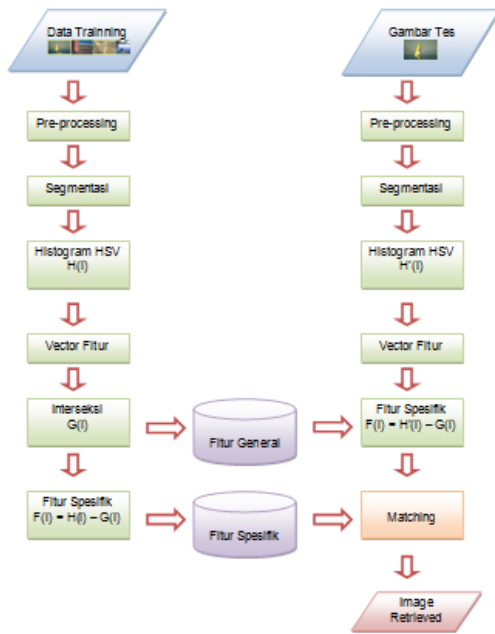


Gambar 2. Gambaran Umum Perancangan Sistem

Gambar 2 menunjukkan sistem bekerja dengan menggunakan *online* kamera untuk memperoleh gambar dan dilakukan proses retrieval.

3.1.1. Perancangan Sistem CBIR

Perancangan perangkat lunak dalam pembuatan sistem terdapat pada gambar 3.

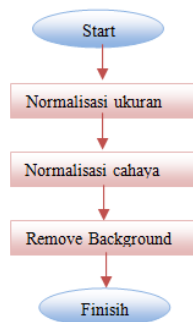


Gambar 3. Blok Diagram Sistem CBIR

Sistem pada gambar 3 adalah memperoleh fitur ciri obyek dengan interseksi histogram.

3.1.2. Pemrosesan awal.

Raw image dari pencarian data ini bervariasi baik kualitas maupun ukuran piksel gambarnya, dikarenakan sumber datanya juga bervariasi. Sehingga perlu dilakukan standarisasi perspektif agar algoritma yang dipakai dalam proses selanjutnya dapat bekerja secara optimal. Pemrosesan awal yang akan dilakukan ada 3 proses yaitu : Normalisasi ukuran (resize 320 x 240), Normalisasi cahaya, Menghilangkan *background*. Diagram alir ada pada gambar 4.



Gambar 4. Diagram alir pemrosesan awal.

3.1.3. Distribusi Warna

Distribusi warna yang digunakan pada program penelitian ini 48 warna crayon.

Tabel 1. Warna 48 Crayon

No.	R	G	B
0	188	138	33
1	42	40	113
2	53	48	55
3	158	63	131
4	26	28	41
5	28	130	44
6	91	48	42
7	2	111	69
8	170	189	193
9	61	48	42
10	0	155	96
11	150	169	186
12	24	83	87
13	1	93	134
14	188	177	1
15	42	74	87
16	64	46	46
17	40	93	68
18	31	127	100
19	193	61	110
20	198	99	145
21	230	201	157
22	209	74	19
23	96	118	53
24	95	164	221
25	33	24	141
26	203	47	22
27	10	88	188
28	76	117	222
29	123	227	216
30	150	141	222
31	62	56	154
32	190	161	197
33	230	143	14
34	14	98	98
35	227	246	252
36	117	121	130
37	179	161	111
38	114	49	47
39	66	34	99
40	171	28	58
41	117	28	58
42	21	154	89
43	149	91	87
44	255	201	101
45	217	215	70
46	145	124	41
47	14	122	214

Tabel 1 merupakan tabel nilai RGB pada 48 warna crayon (x1-x48). Dari tabel 1 maka dapat ditampilkan pada program akan seperti gambar dibawah ini.



Gambar 5. Distribusi Warna Crayon 48 warna

3.1.4. Segmentasi Warna

Segmentasi adalah suatu proses yang digunakan untuk mengelompokkan gambar sesuai dengan obyek gambarnya. Setelah menentukan berapa banyak distribusi warna yang digunakan maka tiap gambar yang ada dilakukan segmentasi warna.

$$e_i = \sqrt{(r_i - r)^2 + (g_i - g)^2 + (b_i - b)^2}$$



Gambar 6. Segmentasi 48 warna

3.1.5. Histogram HSV

HSV merupakan salah satu cara untuk mendefinisikan warna yang didasarkan pada roda warna. *Hue* mengukur sudut sekitar roda warna (merah pada 0 derajat, 120 derajat di hijau, biru, di 240 derajat). *Saturation* yang menunjukkan pada radius roda warna sehingga menunjukkan proporsi antara gelap (pusat) untuk warna ke putih murni (di luar). *Value* menunjukkan nilai kecerahan. *Hue* memiliki nilai antara 0 hingga 360 (derajat), *Saturation* and *Value* berkisar dari 0 hingga 100%.

3.1.6. Interseksi

Histogram interseksi melakukan perhitungan jarak dengan membandingkan dua histogram $h1$ dan $h2$ dan terhadap n bin dengan mengambil nilai interseksi dari kedua histogram tersebut.

$$D_H(h1, h2) = 1 - \frac{\sum_{i=1}^n \min(h1_i, h2_i)}{\sum_{i=1}^n h1_i}$$

3.1.7. Image Matching

Image matching adalah metode yang digunakan dalam Penelitian ini yang bertujuan untuk mengetahui kemiripan dua buah gambar. Pertama, gambar query dan gambar database ditentukan. Gambar query disebut juga dengan gambar acuan atau gambar template adalah suatu gambar yang dijadikan acuan informasi (content) dalam proses pencarian. Sedangkan gambar database atau kumpulan gambar target adalah sekumpulan gambar yang akan digunakan sebagai database gambar.

Kemudian kedua gambar tersebut dilakukan proses ekstraksi ciri warna. Gambar yang mempunyai jarak yang

minimal adalah gambar yang paling mirip dengan gambar template

$$d = |\bar{V}_{(db)} - \bar{V}_{(Q)}|$$

dimana $\bar{V}_{(db)}$ dan $\bar{V}_{(Q)}$ adalah vektor dari gambar database dan gambar query.

4. Hasil dan Pembahasan

Pengujian menggunakan 200 data training gambar, terdiri dari 10 kategori gambar, masing-masing kategori ada 20 gambar (10 gambar asli dari internet dan 10 gambar hasil dari gambar yang diprint setelah itu dicapture menggunakan webcam).































Tabel 2. Data Training

No.	Jenis Gambar
1.	Bunga Sepatu
2.	Bunga Melati
3.	Bunga Matahari
4.	Bunga Mawar merah
5.	Bunga Angrek Ungu
6.	Buah Pear
7.	Tomat
8.	Harimau
9.	Burung berwarna biru
10.	Kuda
11.	Bunga Anggrek

4.1. Pengujian Retrieval pada Bunga Sepatu

Uji coba dilakukan terhadap 5 gambar bunga Sepatu yang sama dengan data training, dan hasil yang ditampilkan adalah 5 gambar dengan urutan nilai prosentasi kemiripan dimulai dari yang terbesar sampai yang terkecil.































Tabel 3. Hasil Uji Coba Bunga Sepatu

Gambar query	Hasil Retrieval					%
	Indeks 1	Indeks 2	Indeks 3	Indeks 4	Indeks 5	
						100
	23029.3	2340.3	34693.4	40491.3	52505.6	
						100
	33970.3	37142.5	37696.8	46248.8	46314.3	
						100
	25219.1	36835.9	36839.2	39901.9	40160.5	
						100
	15447.6	28610.7	32384.1	34331.2	36879.2	
						100
	15447.6	28610.7	32384.1	34331.2	36879.2	
% Rata-rata kemiripan						100

4.2. Pengujian Retrieval pada Bunga Melati

Uji coba dilakukan terhadap 5 gambar bunga Melati yang sama dengan data training, dan hasil yang ditampilkan adalah 5 gambar dengan urutan nilai prosentasi kemiripan dimulai dari yang terbesar sampai yang terkecil.































Tabel 4. Hasil Uji Coba Bunga Melati

Gambar query	Hasil Retrieval					%
	Indeks 1	Indeks 2	Indeks 3	Indeks 4	Indeks 5	
						100
	16891.9	20619.2	23728.5	26517.7	27948.5	
						100
	19476.5	21692.2	24654.5	29405.8	33264.6	
						100
	19042.2	24109.4	27163.3	31333.9	32223.1	
						100
	9962.01	22107.9	25438.5	26805.2	29293.3	
						100
	19563.6	25584.5	33967.1	36130	37748.3	
% Rata-rata kemiripan						100

4.3. Pengujian Retrieval pada Bunga Matahari

Uji coba dilakukan terhadap 5 gambar bunga Matahari yang sama dengan data training, dan hasil yang ditampilkan adalah 5 gambar dengan urutan nilai prosentasi kemiripan dimulai dari yang terbesar sampai yang terkecil.

Tabel 5. Hasil Uji Coba Bunga Matahari

Gambar query	Hasil Retrieval					%
	Indeks 1	Indeks 2	Indeks 3	Indeks 4	Indeks 5	
						100
	24366.9	43878.1	68010.8	69368.7	69893.1	
						100
	46039.1	59354.4	61099.8	63185.8	68060.9	
						100
	33531.8	50025.4	54169.7	56411.7	56820.3	
						100
	37895.9	57531.8	62324.9	69607.1	70493.8	
						100
	45955.1	47876.3	48832.8	49729.2	53563.1	
% Rata-rata kemiripan						100

Pengujian juga dilakukan pada gambar Bunga Mawar, gambar Bunga Anggrek Ungu, gambar Buah Pear, gambar Buah Tomat, gambar Harimau, gambar Burung, dan gambar Kuda.

4.4. Pembahasan

Dari uji coba pada 10 kategori gambar diatas dapat disimpulkan rata-rata persentase hasil retrieval adalah sebagai berikut :

Tabel 6. Kemiripan Obyek

No	Kategori gambar	% Rata – rata kemiripan hasil retrieval
1.	Bunga Sepatu	100
2.	Bunga Melati	100
3.	Bunga Matahari	100
4.	Bunga Mawar Merah	100
5.	Bunga Anggrek Ungu	96
6.	Buah Pear	80
7.	Buah Tomat	100
8.	Harimau	92
9.	Burung Berwarna Biru	96
10.	Kuda	80
% Rata – rata kemiripan		94,4

Berdasar hasil pengujian diperoleh rata-rata kemiripan 94.4%. Kesalahan disebabkan oleh histogram pada data tes yang memiliki kemiripan yang tinggi dengan data tes objek yang lain. Interseksi Histogram HSV cukup baik digunakan untuk memperoleh fitur ciri sebuah obyek. Image Retrieval dengan online kamera dapat dilakukan dan diperoleh hasil yang cukup baik. Pengujian juga dilakukan dengan data gambar diluar data training dengan keberhasilan sebesar 90%.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil pengujian dan analisa pada sistem ini , maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Rata-rata presentase kemiripan gambar dengan uji coba gambar query dari data training sebesar 94.4%
2. Rata-rata presentase kemiripan gambar dengan uji coba gambar query diluar data training sebesar 90%
3. Rata-rata persentase kemiripan berdasarkan dominan warna pada foreground citra untuk sample di atas sebesar 92,2%. Hasil tersebut menunjukkan bahwa ciri warna dominan pada foreground citra sudah dapat membedakan gambar dengan baik.
4. Intensitas cahaya mempengaruhi akurasi pada sistem image retrieval dan diselesaikan dengan proses normalisasi cahaya.
5. Ukuran citra tidak mempengaruhi akurasi pada sistem image retrieval karena dilakukan proses normalisasi ukuran obyek.

5.2. Saran

Berdasar Penelitian yang telah dilakukan masih dapat dikembangkan dengan metode yang lain sehingga didapatkan hasil yang lebih memuaskan. Saran-saran yang dapat diberikan diantaranya adalah:

1. Peningkatan keakuratan hasil identifikasi.
2. Menerapkan metode pengelompokan menggunakan algoritma statistika atau kecerdasan buatan.
3. Menggabungkannya dengan fitur yang lain, untuk mendapatkan hasil yang lebih optimal.

Daftar Pustaka

- [1] Krishnan; Banu M.Sheerin ; Christtiyana ,C Callins.” *Content Based Image Retrieval using Dominant Color Identification Based on Foreground Objects*”
- [2] G.Gordon ; T.Darrell ; M.Harville ; J.Woodfill .“*Background Estimation and removal based on range and color* “
- [3] Ramadijanti, Nana. “*Modul Praktikum Color Histogram* “, PENS -ITS
- [4] Setiawardhana. “ *Identifikasi Kematangan Buah Tomat* “, PENS –ITS
- [5] Susilo, Andriant. “*Web Image Retrieval Untuk Identifikasi Bunga Dengan Pengelompokan Content Menggunakan Ciri Warna dan Bentuk* “ , 2006.